ELSEVIER

Computational Geometry 5 (1995) 155–164

Computational Geometry
Theory and Applications

The expected size of the sphere-of-influence graph

Rex A. Dwyer

Department of Computer Science, North Carolina State University, Raleigh, NC 27695-8206, USA

Communicated by Frances Yao; submitted 21 May 1993; accepted 25 October 1994

Abstract

The *sphere-of-influence graph* of a set of point sites in $\mathbb{R}^d$ is constructed by identifying the nearest neighbor of each site, centering a ball at each site so that its nearest neighbor lies on the boundary, and joining two sites by an edge if and only if their balls intersect. The asymptotic behavior of the expected number of edges of this graph is investigated when the sites are independent and uniformly distributed and their number grows without bound.

1. Introduction

The *sphere-of-influence graph (SIG)* of n point sites $\{P_1, P_2, \ldots, P_n\}$ in $\mathbb{R}^d$ is constructed as follows:

- For each P_i, construct the largest possible empty open ball B_i centered at P_i. (By an "empty" ball, we mean one containing no other sites in its interior.) P_i's nearest neighbor will lie on the boundary of this ball.
- Draw an edge between P_i and P_j if and only if $B_i \cap B_j \neq \emptyset$.

Toussaint [11] proposed this graph as a good *primal sketch* of a dot pattern, suitable for certain low-level vision tasks. He also remarked that it provides a graph-theoretic explanation for the Mueller–Lyer optical illusion, in which equal line segments are perceived to be unequal. Our work is motivated by a potential application to multivariate nonparametric two-sample testing described in Section 5.

Key combinatorial results for sphere-of-influence graphs are Michael and Quint's [9] upper bound of $(5^d - 1.5)n$ edges in d dimensions, and the tighter bound of $17.5n$ for the plane implied by Reifenberg's work [10]. Smith [4] has recently developed an optimal algorithm for constructing SIGs that requires $O(n \log n)$ time in the worst case in any fixed dimension. Other developments are sketched in surveys by Jaromczyk and Toussaint [7] and Michael and Quint [8].

SSDI 0925-7721(94)00025-5

Our task will be to bound the *average* number of edges in the SIG. We will assume that the sites are drawn independently from the uniform distribution on the d-dimensional unit ball. We will see that, for $d \geqslant 7$ and $n \gg d$, the expected number of edges lies between $(0.324)2^d n$ and $(0.677)2^d n$. We will also show that, on average, $O(n)$ time suffices to construct the graph under this model.

In the next section, we will introduce some basic notations of integral geometry and prove some useful geometric and combinatorial lemmata. Section 3 contains the proof of our upper bound and a first estimate of a lower bound. Section 4 sketches a refinement of the lower bound. The final section includes numerical approximations of certain constants, algorithmic considerations, and remarks on applications to nonparametric two-sample testing.

2. Preliminaries

We use some of the usual concepts and notations of integral geometry. $\Gamma(x)$ is the usual extension of the factorial function to the real numbers. B_k represents the unit k-ball, and also its volume. S_k represents both the unit k-dimensional sphere and its surface area. (S_k is the boundary of B_{k+1}.) The usual measure on the unit $(d-1)$-sphere ("surface area") is denoted by σ. Thus, if u ranges over unit vectors in $\mathbb{R}^d$, then

$$\int_{S_{d-1}} d\sigma(u) = S_{d-1} = d \cdot B_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}.$$

(With polar coordinates (r, θ) for $u \in \mathbb{R}^2$, we have $d\sigma(u) = d\theta$; with spherical coordinates in $\mathbb{R}^3$, we have $d\sigma(u) = \sin\phi\, d\phi\, d\theta$.)

The following lemmata will be helpful in later sections.

Lemma 1. *As* $n \to \infty$,

$$\int_0^1 \left(1 - \frac{u}{n}\right)^n u^k du \sim \Gamma(k+1) = k!.$$

Proof. See Whitaker and Watson [12, p. 242]. □

Lemma 2. *For* $m > 0$,

$$\sum_{k \geqslant 0} \frac{1}{m+k}\binom{d}{k} \in \left[\frac{2}{2m+d}2^d, \frac{2m+d}{2m(m+d)}2^d\right].$$

Proof. Exploiting the symmetry of the binomial coefficients, we can write

$$\sum_{k \geqslant 0} \frac{1}{m+k}\binom{d}{k} = \frac{1}{2}\sum_{k \geqslant 0}\left(\frac{1}{m+k} + \frac{1}{m+d-k}\right)\binom{d}{k}$$
$$= \frac{2m+d}{2}\sum_{k \geqslant 0} \frac{1}{(m+k)(m+d-k)}\binom{d}{k}.$$

By differentiating the fraction with respect to k, we find that it is minimized when $k = d/2$ and maximized when $k = 0$. The result follows immediately. □

Since the weight of the binomials is concentrated around $k = d/2$, the value of this sum will tend to the lower limit if $m = O(d)$ and $d \to \infty$. (This could be proved in detail using the DeMoivre–Laplace Theorem.)

Lemma 3. *Let s_1 and s_2 be two spheres of radius r_1 and r_2 with centers at distance t. Let*

$$g(r_1, r_2, t) = \frac{\sigma(s_1 \setminus s_2)}{\sigma(s_1)} \cdot \frac{\sigma(s_2 \setminus s_1)}{\sigma(s_2)}.$$

If neither sphere contains the other's center, then

$$4/9 \leqslant g(r_1, r_2, t) \leqslant 1 \quad \text{for } d = 2,$$

and

$$1/2 \leqslant g(r_1, r_2, t) \leqslant 1 \quad \text{for } d > 2.$$

Proof. Since neither sphere contains the other's center, $r_1 \leqslant t \leqslant r_1 + r_2$. The upper bound is achieved whenever $t = r_1 + r_2$. The lower bound is achieved when $t = r_1$, but it is necessary to determine the relationship between r_1 and r_2 in this case.

Without loss of generality, let us assume that $r_2 \geqslant r_1$. A tedious analysis will show that g increases to a maximum then decreases as r_2 increases from 0 to r_1, so the minimum is achieved either when $r_2 = 0$ or when $r_2 = r_1$. In any dimension, $g \to 1/2$ as $r_2 \to 0$. When $r_1 = r_2$, $g = 4/9$ for $d = 2$, $g = 9/16$ for $d = 3$, etc., with g increasing monotonically with d. □

3. Derivation of bounds

P_1 and P_2 can be sphere-of-influence neighbors under four distinct circumstances:

- when P_1 and P_2 are each other's nearest neighbors,
- when P_1 is P_2's nearest neighbor or vice versa, but not both,
- when P_1 and P_2 share a common nearest neighbor P_3, and their spheres of influence intersect, or
- when P_1 and P_2 have distinct nearest neighbors P_3 and P_4, and their spheres of influence intersect.

The edges of the first two types are precisely the edges of the nearest-neighbor graph, which has been thoroughly studied. The expected number of nearest-neighbor edges is $\mathbb{E}NN_n \sim 0.68n$ for $d = 2$, and the constant of dimensionality tends quickly to the limit 0.75 as d grows. (This constant is

$$\frac{3\sqrt{\pi}\,\Gamma((d+1)/2) - 4\Gamma((d+2)/2)J}{4\sqrt{\pi}\,\Gamma((d+1)/2) - 4\Gamma((d+2)/2)J}, \qquad \text{where } J = \int_0^{\pi/3} (\sin\theta)^d d\theta, \tag{3.1}$$

to be precise.) When the third situation obtains, we say that (P_1, P_2, P_3) is a *SIG triple*, and in the fourth case, we call (P_1, P_2, P_3, P_4) a *SIG quadruple*. We write $\mathbb{E}T_n$ and $\mathbb{E}Q_n$ for the expected number of triples and quadruples respectively. Thus, the expected size of the sphere-of-influence graph satisfies

$$\mathbb{E}SIG_n = \mathbb{E}NN_n + \mathbb{E}T_n + \mathbb{E}Q_n.$$

Not surprisingly, $\mathbb{E}Q_n$ is the most significant term, so we analyze it first.

We call (P_1, P_2, P_3, P_4) a *SIG quadruple* if and only if

- P_1P_3 is the radius of an empty ball with P_1 at its center,
- P_2P_4 is the radius of an empty ball with P_2 at its center,
- these two balls intersect, and
- $|P_1P_3| \geqslant |P_2P_4|$.

The probability that any SIG edge possesses more than one SIG quadruple is nil, since this would imply that two or more sites lie at exactly the same distance from either P_1 or P_2. Since the four sites play distinct rôles in the SIG quadruple, the number of candidate quadruples is $n^4 = n(n-1)(n-2)(n-3) \sim n^4$. Since the sites are i.i.d., we can determine the number of SIG quadruples by computing the probability that the first four sites form a SIG quadruple. Let G be the probability content of the union of the two balls defined by a candidate SIG quadruple. Then the expected number of SIG quadruples is

$$\mathbb{E}Q_n = n^4 B_d^{-4} \int (1-G)^{n-4}\, dP_1\, dP_2\, dP_3\, dP_4,$$

where integration is over all configurations defining intersecting balls.

To isolate the most significant aspects of the configuration of the four points, we rewrite in terms of spherical coordinates. The origin of the spherical coordinate systems is not fixed, however: P_1 is the origin for P_2 and P_3, and P_2 is the origin for P_4. In this system, (q, v_1), (t, v_2), (r_1, u_1), and (r_2, u_2) are (radius, unit vector) pairs such that $P_1 = qv_1$; $P_2 = P_1 + tv_2$; $P_3 = P_1 + r_1u_1$; $P_4 = P_2 + r_2u_2$; $r_1 \geqslant r_2 \geqslant 0$; $t \geqslant 0$; and $q \geqslant 0$. It is not too difficult to show that we do not err in neglecting configurations for which $q + 3r_1 > 1$; this assumption implies that the entire configuration, including the two balls, lies entirely inside the unit d-ball. Thus we have

$$\begin{aligned}
\mathbb{E}Q_n &= n^4 B_d^{-4} \int (1-G)^{n-4} (qtr_1r_2)^{d-1}\, dq\, d\sigma(v_1)\, dt\, d\sigma(v_2)\, dr_1 d\sigma(u_1) \\
&\quad \times dr_2 d\sigma(u_2) \\
&= n^4 B_d^{-4} S_{d-1}^4 \int_0^{1/3} \int_0^{r_1} \int_{r_1}^{r_1+r_2} \int_0^{1-3r_1} (1-G)^{n-4} (qtr_1r_2)^{d-1} \\
&\quad \times g(t, r_1, r_2)\, dq\, dt\, dr_2 dr_1 \\
&= n^4 d^3 \int_0^{1/3} \int_0^{r_1} \int_{r_1}^{r_1+r_2} (1-G)^{n-4} (tr_1r_2)^{d-1} (1-3r_1)^d \\
&\quad \times g(t, r_1, r_2)\, dt\, dr_2 dr_1. \qquad (3.2)
\end{aligned}$$

While $g(t, r_1, r_2)$ depends in a rather complicated way on t, r_1, and r_2, according to Lemma 3 and the Mean Value Theorem a constant $\bar{g} \in [1/2, 1]$ exists for which

$$\begin{aligned}\mathbb{E}Q_n &= n^4 d^3 \bar{g} \int_0^{1/3} \int_0^{r_1} \int_{r_1}^{r_1+r_2} (1-G)^{n-4} (t r_1 r_2)^{d-1} (1-3r_1)^d \, dt \, dr_2 \, dr_1 \\ &= n^4 d^2 \bar{g} \int_0^{1/3} \int_0^{r_1} (1-G)^{n-4} (r_1 r_2)^{d-1} (1-3r_1)^d \big((r_1+r_2)^d - r_1^d\big) dr_2 \, dr_1.\end{aligned}$$

Substituting $r_2 = \alpha r_1$ gives

$$\mathbb{E}Q_n = n^4 d^2 \bar{g} \int_0^1 \int_0^{1/3} (1-G)^{n-4} \alpha^{d-1} r_1^{3d-1} (1-3r_1)^d \big((1+\alpha)^d - 1\big) dr_1 \, d\alpha. \tag{3.3}$$

We estimate

$$(1+\alpha^d/2) r_1^d \leqslant G \leqslant (1+\alpha^d) r_1^d. \tag{3.4}$$

Setting

$$u = n(1+\alpha^d) r_1^d \qquad \text{and} \qquad \frac{du}{n(1+\alpha^d) d} = r_1^{d-1} dr_1,$$

we obtain the lower bound

$$\begin{aligned}\mathbb{E}Q_n &\geqslant n^4 d^2 \bar{g} \int_0^1 \int_0^{n(1+\alpha^d)/3^d} \left(1 - \frac{u}{n}\right)^{n-4} \alpha^{d-1} \left(\frac{u}{n(1+\alpha^d)}\right)^2 \\ &\quad \times \left(1 - \frac{3u}{n(1+\alpha^d)}\right)^d \big((1+\alpha)^d - 1\big) \frac{du}{n(1+\alpha^d) d} d\alpha \\ &\sim n d \bar{g} \left(\int_0^\infty \left(1 - \frac{u}{n}\right)^{n-4} u^2 du\right) \left(\int_0^1 \alpha^{d-1} (1+\alpha^d)^{-3} \big((1+\alpha)^d - 1\big) d\alpha\right) \\ &= n(2\bar{g}d) \int_0^1 \alpha^{d-1} (1+\alpha^d)^{-3} \big((1+\alpha)^d - 1\big) d\alpha.\end{aligned} \tag{3.5}$$

For a lower bound, we expand a power series about $\alpha = 1$ to find that $(1+\alpha^d)^{-3} \geqslant (8 - 9\alpha^d + 3\alpha^{2d})/16$ for $0 \leqslant \alpha \leqslant 1$. Therefore, by expanding the binomial $(1+\alpha)^d$ and applying Lemma 2, we find

$$\begin{aligned}&\int_0^1 \alpha^{d-1} (1+\alpha^d)^{-3} \big((1+\alpha)^d - 1\big) d\alpha \\ &\geqslant \frac{1}{16} \sum_{k \geqslant 1} \binom{d}{k} \int_0^1 (8\alpha^{d+k-1} - 9\alpha^{2d+k-1} + 3\alpha^{3d+k-1}) d\alpha \\ &\geqslant \frac{1}{16} \left(\sum_{k \geqslant 0} \frac{8}{d+k} \binom{d}{k} - \sum_{k \geqslant 0} \frac{9}{2d+k} \binom{d}{k} + \sum_{k \geqslant 0} \frac{3}{3d+k} \binom{d}{k}\right) - \frac{9}{32d} \\ &\geqslant \frac{17 \cdot 2^d}{105d} - \frac{9}{32d} \approx (0.162) 2^d / d\end{aligned}$$

Substituting this approximation and $\bar{g} \geqslant 1/2$ into 3.5, we obtain the lower bound

$$(1 + o(1))\mathbb{E}Q_n \geqslant (0.162)2^d n.$$

From (3.3) and (3.4) with $u = (1 + \alpha^d/2)r_1^d$, we obtain the upper bound

$$\mathbb{E}Q_n \leqslant (1 + o(1))n(2\bar{g}d)\int_0^1 \alpha^{d-1}(1 + \alpha^d/2)^{-3}\left((1 + \alpha)^d - 1\right)d\alpha. \tag{3.6}$$

Since $(1 - \alpha^d/2)^{-3} \leqslant 1 - \frac{10}{9}\alpha^d + \frac{11}{27}\alpha^{2d}$ for $0 \leqslant \alpha \leqslant 1$, the integral lies in the interval

$$\left[\frac{64 \cdot 2^d}{189d} - \frac{47}{81d}, \frac{263 \cdot 2^d}{648d} - \frac{47}{81d}\right]$$

and tends to the lower limit. Since $\bar{g} \leqslant 1$, we obtain $\mathbb{E}Q_n \leqslant (0.812)2^d n$ for all d, and for large d,

$$(0.162)2^d n \leqslant (1 + o(1))\mathbb{E}Q_n \leqslant (0.677)2^d n.$$

Only the SIG triples remain. We assume that P_3 is the common nearest neighbor of P_1 and P_2 and that $|P_1P_3| \geqslant |P_2P_3|$. We have

$$\mathbb{E}T_n \sim n^3 B_d^{-3}\int (1 - G)^{n-2}\,dP_1\,dP_2\,dP_3,$$

where integration is again over all configurations with overlapping circles. We rewrite in generalized spherical coordinates so that $P_1 = P_3 + r_1u_1$; $P_2 = P_3 + r_2u_2$; and $P_3 = qv$ for unit vectors u_1, u_2, and v and real numbers $q, r_1, r_2 \geqslant 0$. Again, we restrict our attention to $q + 2r_1 \leqslant 1$ to guarantee that both balls are contained in the unit ball. We write $h(r_1, r_2)$ for the proportion of the surface area of a sphere of radius r_2 that lies outside a sphere of radius $r_1 \geqslant r_2$ when the smaller sphere's center lies on the larger's surface. It is not hard to show that $1/2 \leqslant h \leqslant 1$. In fact, the maximum is achieved when $r_2 = r_1$ and

$$h = h_{\max} = 1 - \frac{S_{d-2}}{S_{d-1}}\int_{1/2}^1 (1 - t^2)^{(d-3)/2}\,dt. \tag{3.7}$$

In light of all this, we know that for some $\hbar \in [1/2, h_{\max}]$

$$\begin{aligned}\mathbb{E}T_n &\sim n^3 B_d^{-3}\int (1 - G)^{n-2}(qr_1r_2)^{d-1}\,d\sigma(v)\,d\sigma(u_1)\,d\sigma(u_2)\,dq\,dr_2\,dr_1\\ &= n^3 d^2\int_0^{1/2}\int_0^{r_1}(1 - G)^{n-2}(r_1r_2)^{d-1}h(r_1, r_2)\,dr_2\,dr_1\\ &= n^3 d^2\hbar\int_0^1\int_0^{1/2}(1 - G)^{n-2}\alpha^{d-1}r_1^{2d-1}\,dr_1\,d\alpha.\end{aligned}$$

Applying (3.4), we obtain the upper bound

$$ndh_{\max}\left(\int_0^\infty\left(1 - \frac{u}{n}\right)^{n-3}u\,du\right)\left(\int_0^1 \alpha^{d-1}(1 + \alpha^d/2)^{-2}\,d\alpha\right) = \frac{2h_{\max}n}{3}$$

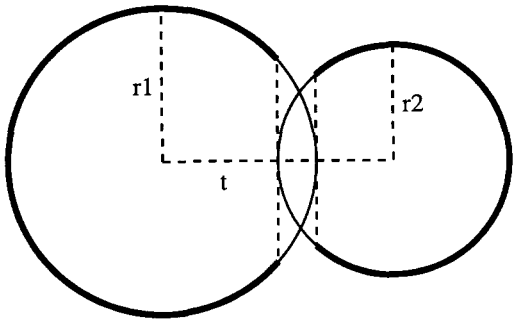

Fig. 1. Lemma 4.

and the lower bound

$$nd(1/2)\left(\int_0^\infty \left(1-\frac{u}{n}\right)^{n-3} u\,du\right)\left(\int_0^1 \alpha^{d-1}(1+\alpha^d)^{-2}\,d\alpha\right) = n/4,$$

i.e., $n/4 \leqslant \mathbb{E}T_n \leqslant 2n/3$ for large d.

4. An improved lower bound

For $d \geqslant 3$, we can improve our lower bound on the number of SIG quadruples by a factor of 2 if we use a better lower bound on $g(r_1, r_2, t)$. This improvement is sketched below. The details are straightforward but quite tedious. They are easy to verify with any good software package for computer algebra.

Lemma 4. *If $d \geqslant 3$, then*

$$g(r_1, r_2, t) \geqslant \frac{r_2+t-r_1}{2r_2}\cdot\frac{r_1+t-r_2}{2r_1} = \frac{t^2-(r_1-r_2)^2}{4r_1r_2}.$$

Proof. The two factors estimate the surface areas illustrated in Fig. 1, which clearly underestimate $g(r_1, r_2, t)$. For $d=3$, it is well known that the surface area of a spherical cap is proportional to its height, thus equality is achieved in the lemma. It is not difficult to show that the inequality holds strictly for $d>3$. □

Picking up the derivation of $\mathbb{E}Q_n$ at (3.2) and ignoring the insignificant factor $(1-3r_1)^d$, we have (for large enough d)

$$\mathbb{E}Q_n \geqslant \frac{n^4d^3}{4}\int_0^{1/3}\int_0^{r_1}\int_{r_1}^{r_1+r_2}(1-G)^n(r_1r_2)^{d-2}t^{d-1}\left(t^2-(r_1-r_2)^2\right)dt\,dr_2\,dr_1$$

$$\sim \frac{d^2n}{2}\left(\int_0^1 \frac{\alpha^{d-2}}{(1+\alpha^d)^3}\left(\frac{(1+\alpha)^{d+2}-1}{d+2}-\frac{(1+\alpha)^d-1}{d}\right)d\alpha\right). \qquad (4.1)$$

Table 1
Numerical estimates of $\lim \mathbb{E}(SIG_n / n)$

d	2	3	4	5	6	7	8	9	10	20	$d \to \infty$
lower	1.39	3.07	5.81	11.4	22.7	45.6	91.9	185.	373.	396293.	$(0.324)2^d$
upper	2.88	5.62	11.0	21.7	43.0	85.4	170.	339.	676.	685713.	$(0.677)2^d$

The integral in α can be bounded by estimating $(1 + \alpha^d)^{-3} \geqslant (8 - 9\alpha^d + 3\alpha^{2d})/16$ and applying Lemma 2 as before; this calculation is straightforward but quite lengthy. The result is the lower bound

$$\mathbb{E}Q_n \geqslant \frac{34}{105} 2^d n \approx (0.324)2^d n$$

for large enough d.

5. Conclusions

Table 1 gives numerical estimates of the upper and lower bounds on $\lim_{n \to \infty} \mathbb{E}(SIG_n/n)$ for small dimensions. For these estimates, the values of the integrals in (3.1), (3.6), (3.7), and (4.1) were computed exactly. It is apparent from the table that the bounds

$$(0.324)2^d \leqslant \lim_{n \to \infty} \mathbb{E}(SIG_n/n) \leqslant (0.677)2^d$$

hold for $d \geqslant 7$.

The gap between the upper and lower bounds comes from three sources: the power-series approximations used to estimate the integrals in (3.6) and (4.1), the estimate of g provided by Lemmata 3 and 4, and the estimate of G in (3.4). Using more terms in the power series can reduce the gap only slightly. A substantial improvement is to be had only by employing better estimates of g and G. Better estimates are easily obtained, but the resulting integrals are difficult to evaluate.

An important question is whether these results apply to other distributions. In fact, they carry over to any distribution absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$. If a SIG edge P_iP_j has length t, then the empty balls B_i and B_j are contained in a ball with radius $3t/2$ centered at the midpoint of the edge. Thus, in the limit, the empty region "causing" the edge's presence in the graph lies in a small neighborhood of the edge in which the distribution may be regarded as uniform. This is the essential ingredient in Devroye's analyses of other geometric proximity graphs [3]. This ingredient is lacking in, for example, the Voronoi dual graph, in which the empty ball causing the edge may extend arbitrarily far beyond the midpoint of the edge.

Our results imply that the sphere-of-influence graph can be constructed in $O(n)$ time on average for uniform points in any fixed dimension d. One algorithm begins by using Bentley, Weide, and Yao's "spiral search" method [2] to identify the nearest neighbor of each point. This algorithm divides the space into $\Theta(n)$ hypercubic cells, assigns the

sites to cells, and finds the nearest neighbor of each site by searching the cells layer by layer in increasing distance from the cell containing the site. It is easily proved that, on average, $O(1)$ cells are searched for each site. Next, each site is inserted into every cell intersecting the boundary of its nearest-neighbor ball. Since most balls are small, on average, each site is inserted into a constant number of cells, and a constant number of sites is inserted into each cell. Finally, each cell is examined. Every pair of sites assigned to the cell is checked for possible intersection of nearest-neighbor balls. In two dimensions, every edge will be detected at most twice, since two circles intersect in two points in at most two different cells. In higher dimensions, the number of times an edge will be detected is potentially large ($\Omega(n^{1-2/d})$) in the worst case, but constant on average, the constant growing exponentially with the dimension. Other algorithmic approaches are necessary when the distribution is not uniform.

A promising application of sphere-of-influence graphs is as a basis for nonparametric multivariate two-sample testing. We imagine a set of red points and a set of blue points, and we wish to test the hypothesis that the two populations were drawn from the same distribution. One approach is to construct a graph in which the presence of an edge indicates the geometric proximity of its endpoints. If too small a fraction of the edges join red points to blue points, we reject the null hypothesis [6]. While the Voronoi dual graph has proven to be useful in this context [5], it may be unsatisfactory when some components of the data are redundant, and the data lies on a manifold of smaller dimension in $\mathbb{R}^d$. In this case, the Voronoi dual graph may include many edges that do not reflect proximity on the data manifold. If the size of the data set is large enough, the sphere-of-influence graph includes very few edges that do not roughly follow the contours of the underlying manifold. As an extreme case, consider two samples lying on distinct orthogonal great circles of a sphere. The Voronoi dual will be a complete graph, since the sphere itself is empty. The sphere-of-influence graph, on the other hand, will contain edges between the two samples only in the neighborhoods of the intersections of the great circles. Banerjia [1] has given a preliminary positive report on the utility of the sphere-of-influence graph for two-sample testing.

Note added in proof. Lemma 4 can be improved further to

$$g(r_1, r_2, t) \geqslant \frac{r_2 + t - r_1}{2r_2}.$$

Table 1 has been updated to reflect this result. The $d \to \infty$ entry is unaffected.

References

[1] S. Banerjia, Proximity graphs and multivariate two-sample testing, Master's thesis, Computer Science Dept., North Carolina State U., 1993.

[2] J.L. Bentley, B.W. Weide and A.C. Yao, Optimal expected-time algorithms for closest-point problems. ACM Trans. Math. Software 6 (1980) 563–580.

[3] L.P. Devroye, The expected size of some graphs in computational geometry, Comput. Math. Appl. 15 (1988) 53–64.
[4] R.A. Dwyer and F.A. Smith, An optimal algorithm for the sphere-of-influence graph in higher dimensions, manuscript, 1994.
[5] R.A. Dwyer and M.B. Squire, A multivariate two-sample test using the Voronoi diagram, Technical Report TR-93-21, Computer Science Dept., North Carolina State U., 1993.
[6] J.H. Friedman and L.C. Rafsky, Multivariate generalizations of the Walds–Wolfowitz and Smirnov two-sample tests, Ann. Statistics 7 (1979) 697–717.
[7] J.W. Jaromczyk and G.T. Toussaint, Relative neighborhood graphs and their relatives. *Proc. IEEE* 80 (1992) 1502–1517.
[8] T.S. Michael and T. Quint, Sphere-of-influence graphs: a survey, manuscript, 1994.
[9] T.S. Michael and T. Quint, Sphere-of-influence graphs: edge density and clique size, manuscript, 1994.
[10] E.R. Reifenberg, A problem on circles, Math. Gaz. 32 (1948) 290–292.
[11] G.T. Toussaint, A graph-theoretical primal sketch, in: G.T. Toussaint, ed., Computational Morphology (Elsevier, Amsterdam, 1988) 229–260.
[12] E.T. Whittaker and G.N. Watson, A Course of Modern Analysis (Cambridge Univ. Press, Cambridge, 4th Ed., 1927).